# Search for Higgs Bosons Produced in Association with b-Quarks at CDF

Thomas Wright
University of Michigan

Fermilab Joint Experimental-Theoretical Seminar
July 8, 2011

# Outline

- Overview of the Higgs sector of the MSSM
- Data samples and the CDF detector
- Search for associated Higgs production in the 3b channel
  - Background templates and fitting
  - MSSM interpretation, scenario dependence, and width effects
- Conclusion and outlook

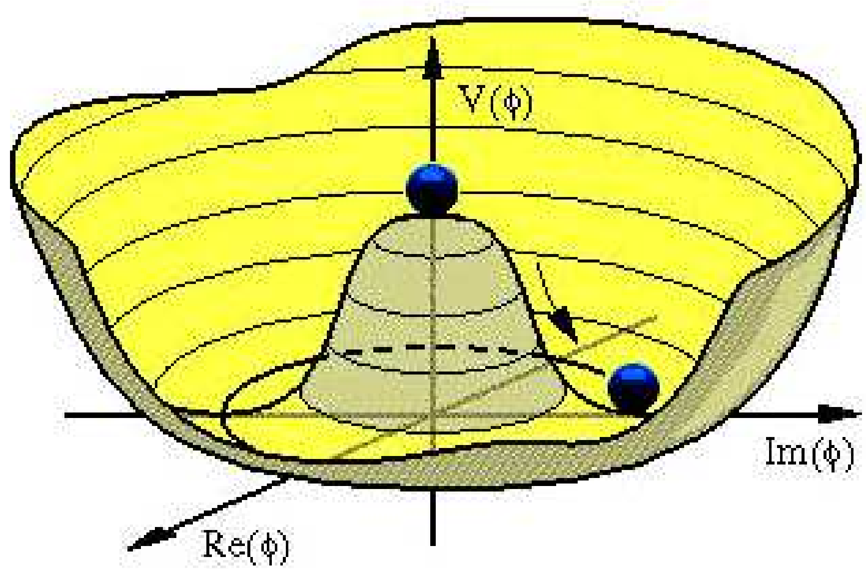
- Results presented today have been submitted to PRD
  - Available at arXiv:1106.4782
  - http://www-cdf.fnal.gov/physics/new/hdg/Results.html

# Why Do We Expect a Higgs?

- Standard Model fermion/boson interactions specified by $SU(2)_L \times U(1)_Y$ symmetry

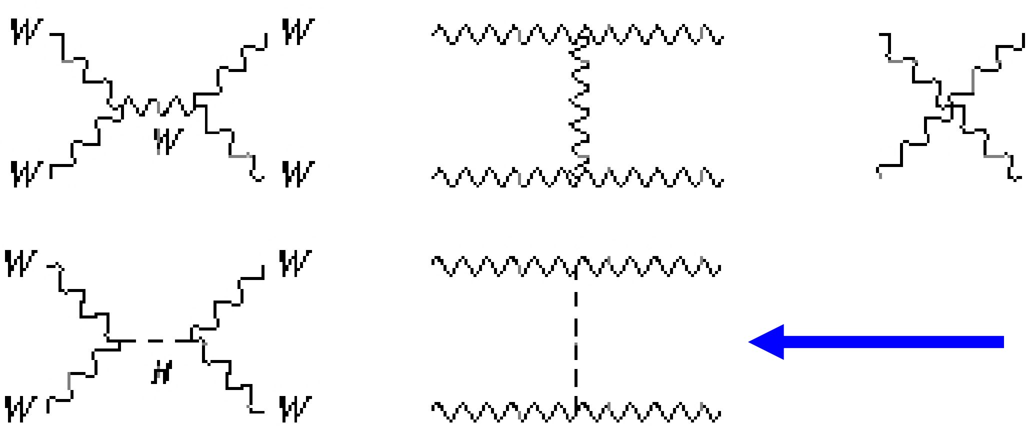
- But, fermion and boson mass terms are forbidden by same symmetry!

- Introduce pair of complex scalar fields with a particular shape, and interactions with fermions/bosons
- All compatible with $SU(2)_L \times U(1)_Y$
- Then break the symmetry




- Excitations away from mimimum
  - Physical Higgs scalar
- Original interactions split
  - VEV terms – masses!
  - Interactions with physical Higgs strength proportional to mass

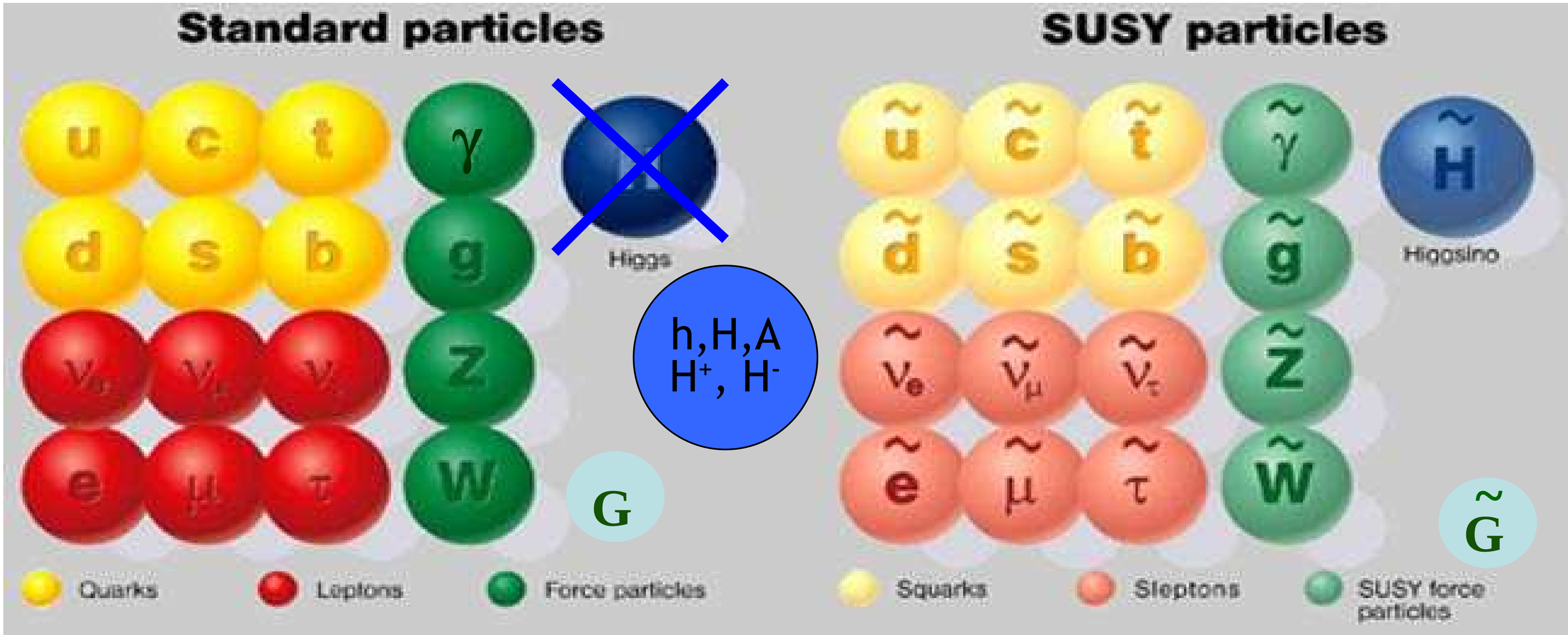
WW scattering in the SM




$\sigma \sim E^2 \rightarrow$ violates unitarity above 1.2 TeV

Physical Higgs has just the right couplings to cancel divergence!

# Supersymmetry




- SM particles have supersymmetric partners: differ by 1/2 unit in spin
- SUSY has many attractive properties
  - Cancellation of Higgs mass divergence, coupling unification, etc
  - Lightest neutralino is a dark matter candidate

- Requires larger Higgs sector than the single scalar of the SM
  - Simplest case: Minimal Supersymmetric Standard Model (MSSM)

# Higgs in the MSSM

- Instead of one scalar, get five:
  - Three neutral: h, H, A : (generically "$\phi$")
  - Two charged: $H^+$, $H^-$
- Separate couplings for up-type and down-type fermions

- Properties of the Higgs sector largely determined by two parameters:
  - $m_A$ : mass of pseudoscalar
  - $\tan\beta$ : ratio of down-type to up-type couplings

- Typically, $m_h < m_A < m_H$ , and $m_{H\pm} \sim m_A$

- For $\tan\beta$ near 1, h is SM-like and light - LEP-II limits apply

- Larger $\tan\beta$ shows more interesting behavior
  - A becomes degenerate with h or H (mass, couplings, etc)
  - Other decouples, SM-like, mass around 120 GeV
  - A + h/H production controlled by $\tan\beta$

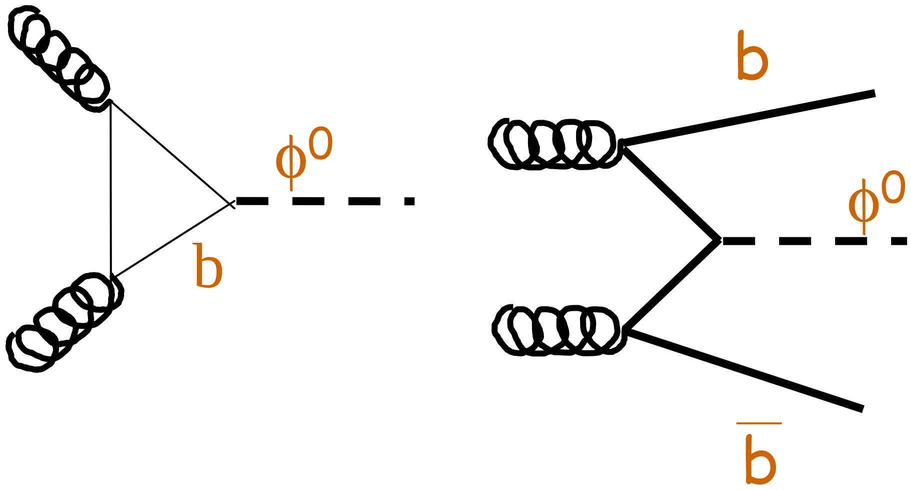
- In the Standard Model, Higgs cross section is fixed – no free parameters
- In MSSM, production of A/h/H depends on $\tan\beta$ – range of possiblity
  - For the right value of $\tan\beta$, could already have discovery potential

# Higgs at High tanβ

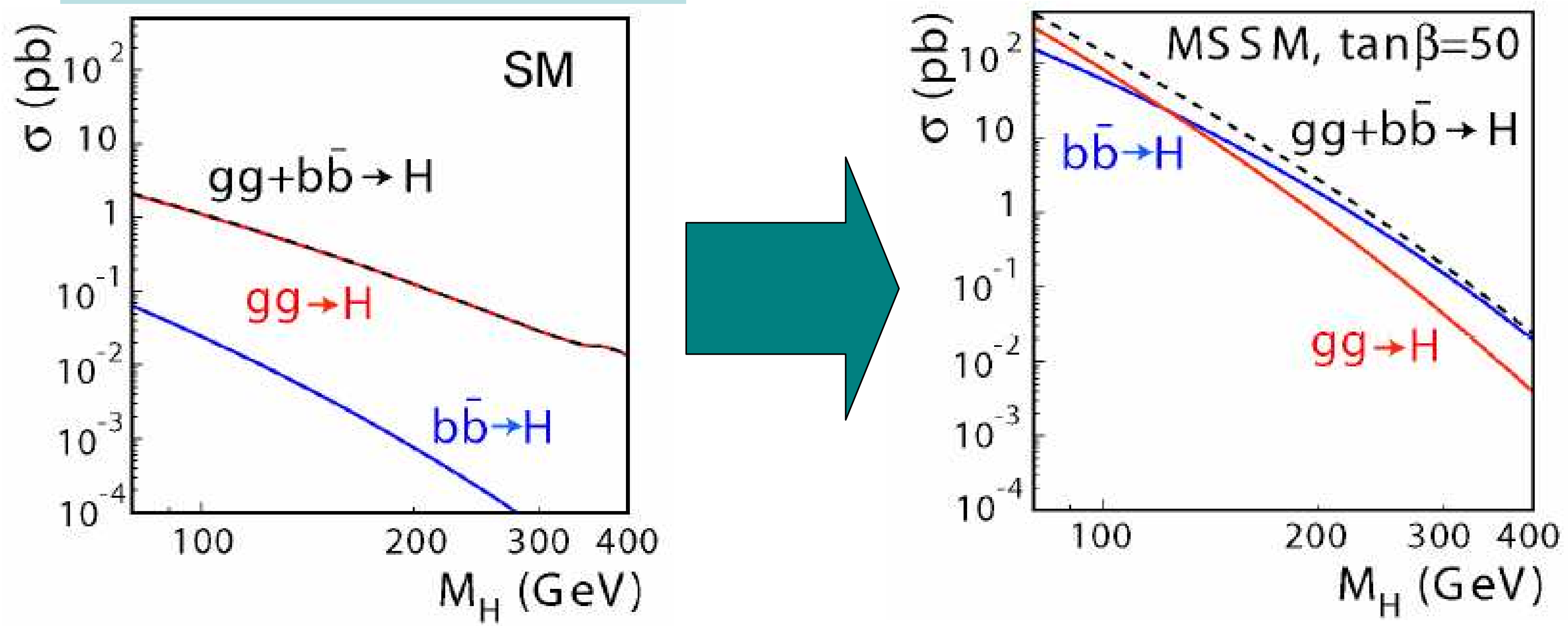
Processes involving bottom quarks (down-type) enhanced by $\tan^2\beta$

Boost from femtobarns to picobarns

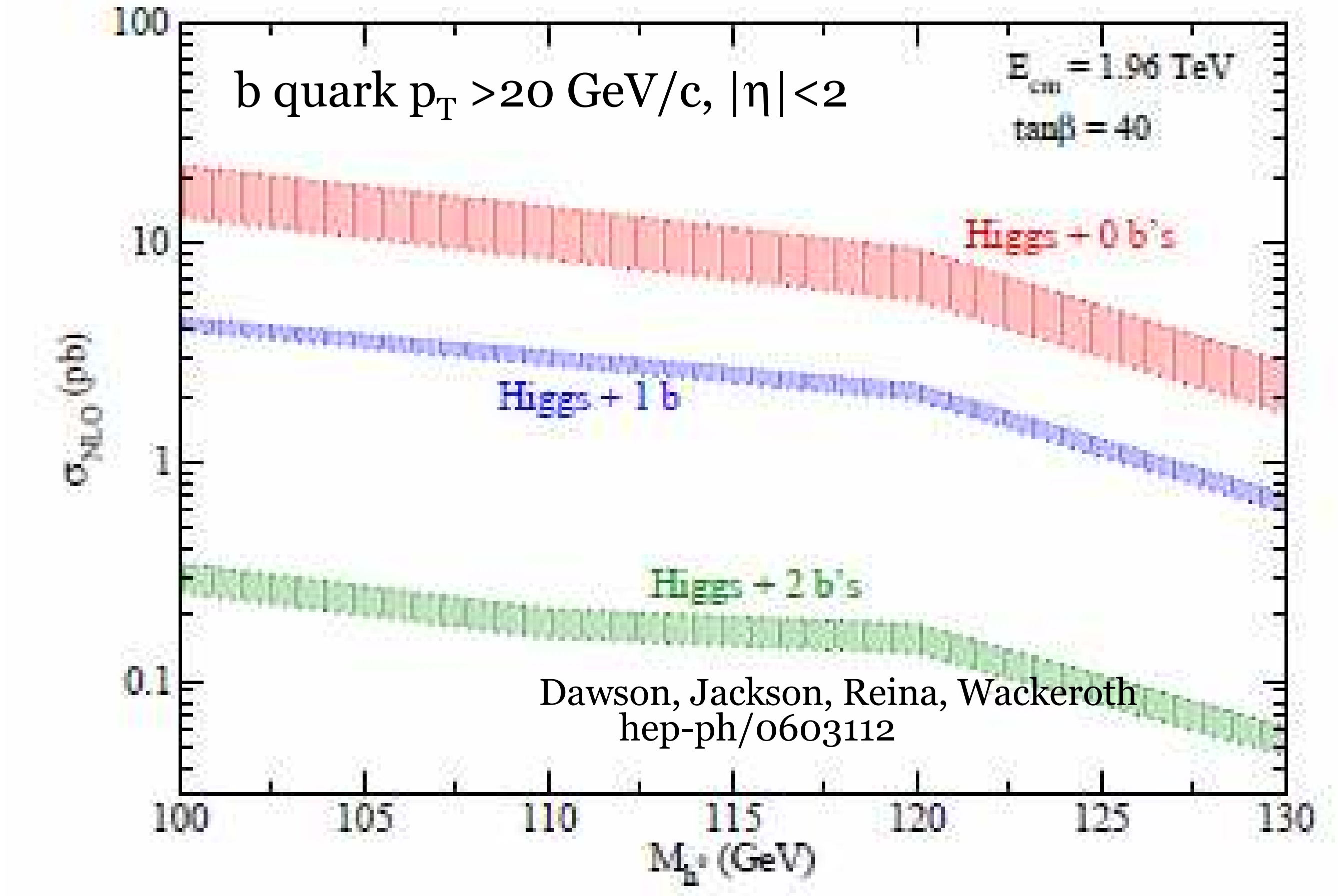
Could be observable at Tevatron!

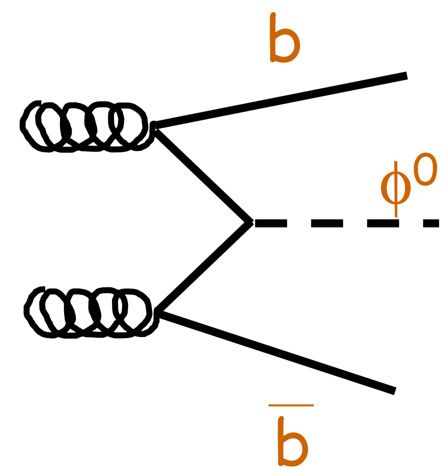



At large tanβ, decays into $b\bar{b}$ (90%) and $\tau\tau$ (10%) dominate



# The 3b Channel



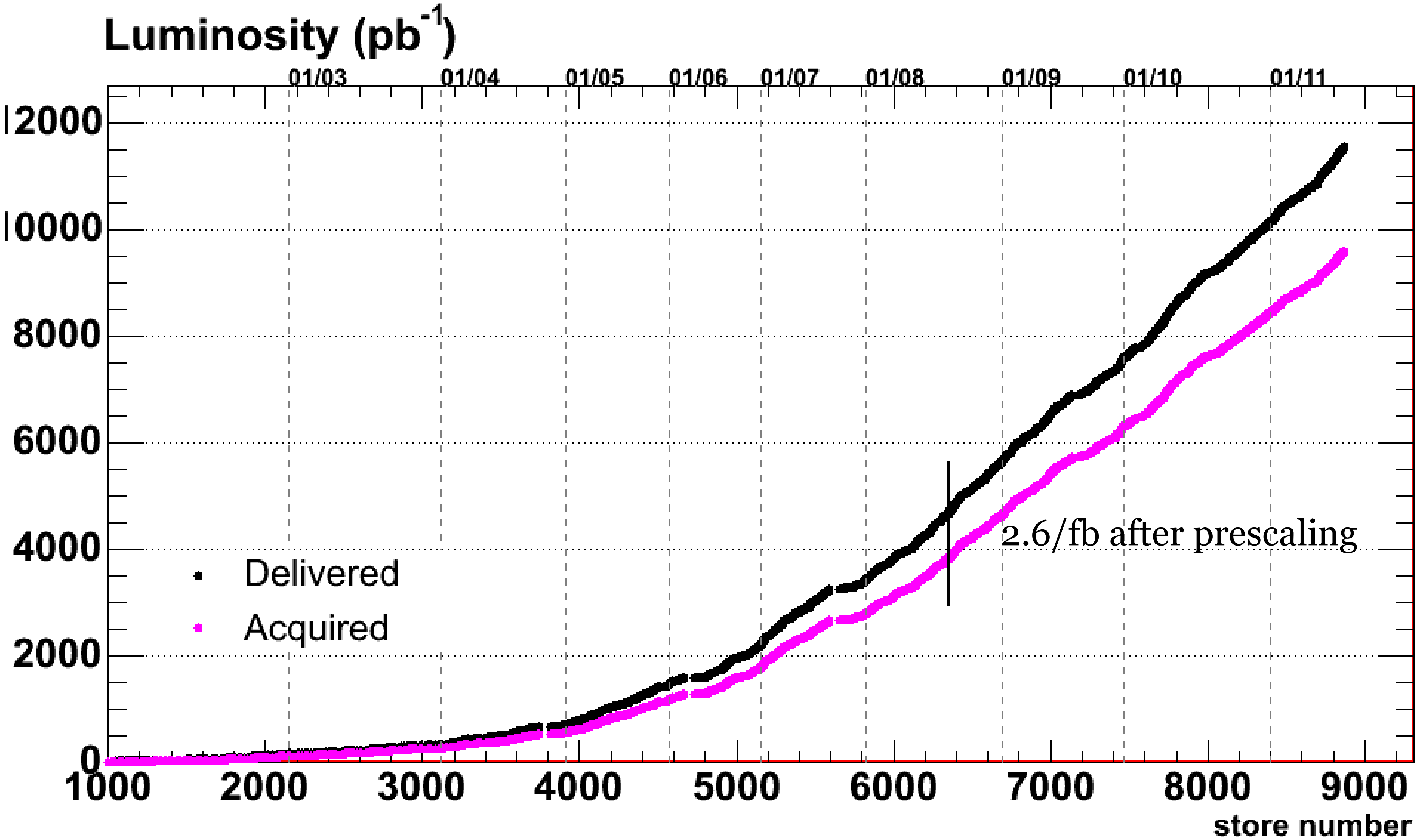




- Search for the bbφ→bbbb process
- Less cross section when requiring both b's to be high-$p_T$

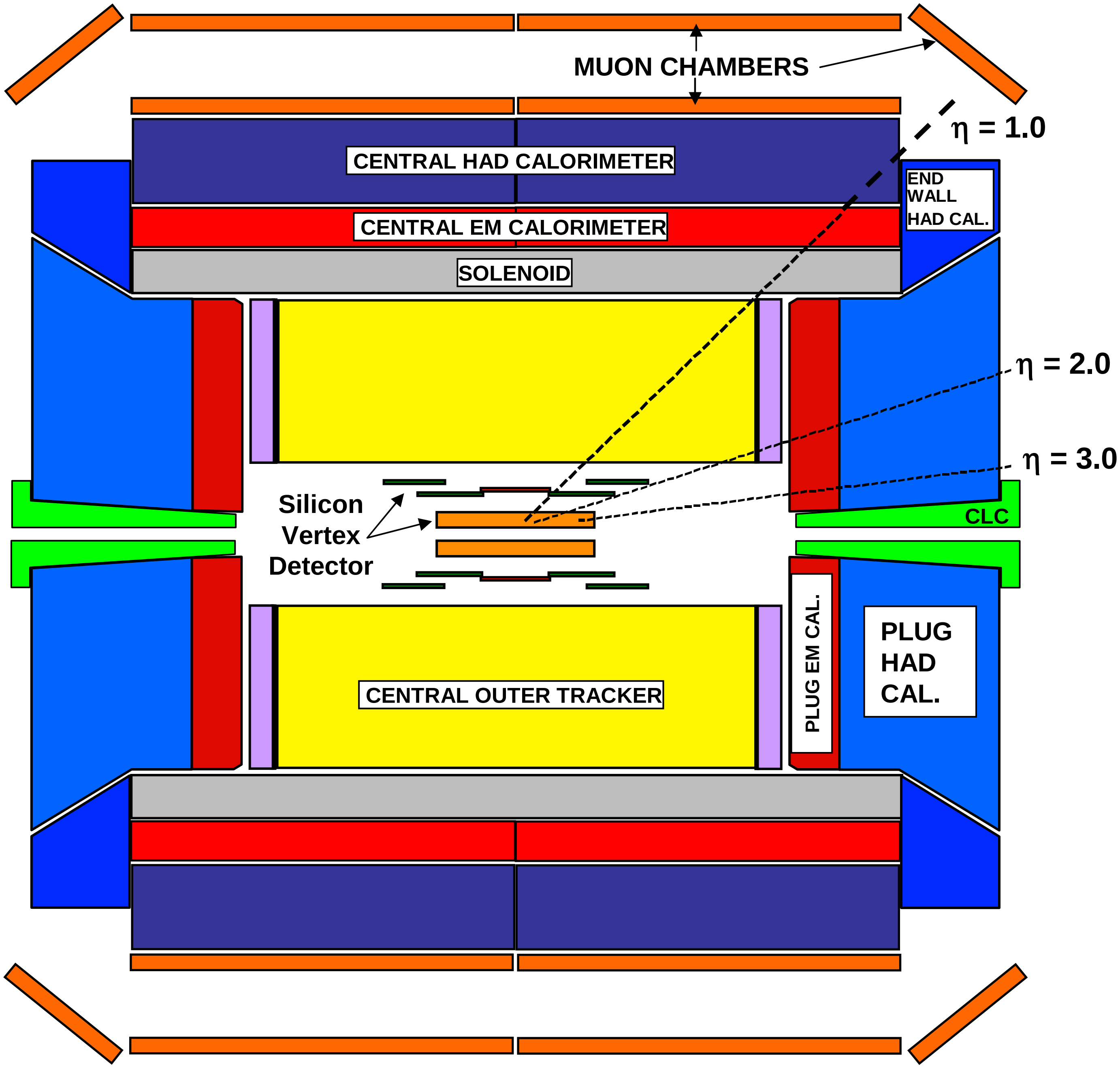
- Look for the Higgs + 1b case
  - Three b-jets with two forming a mass peak

# Data Sample

- Similar displaced-track triggering (SVT) as CDF B-physics program
- Results today use all data up to trigger change (2.6/fb)
- Continue to collect data on new triggers – will analyze full Run II sample



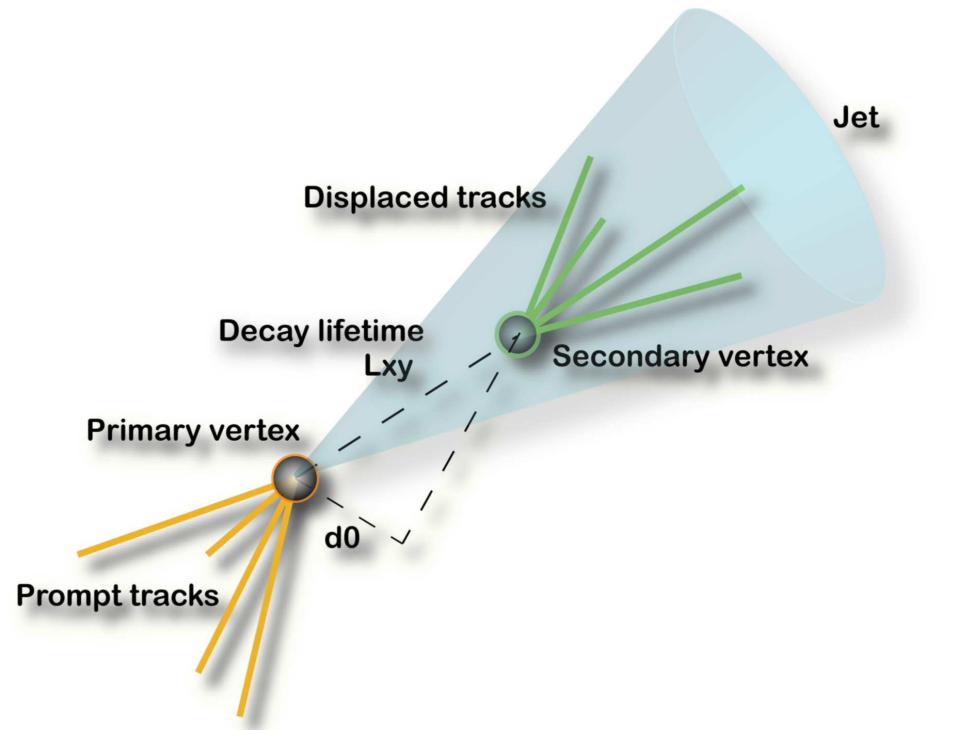
# The CDF II Detector

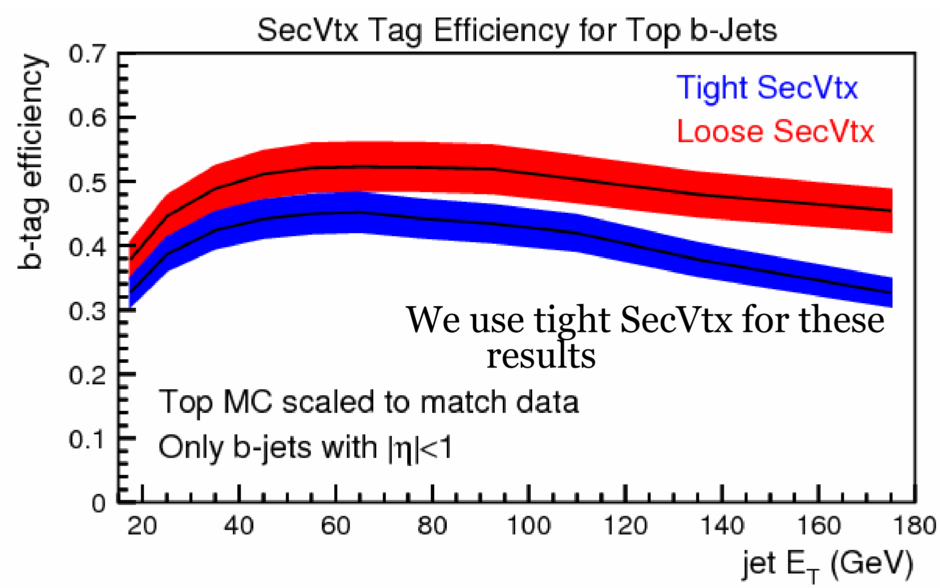


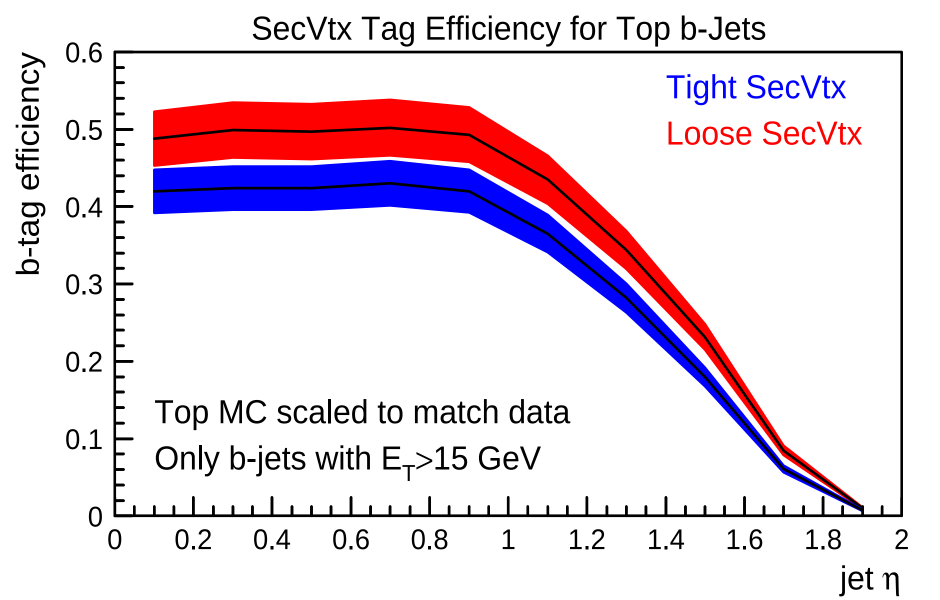
# B-Jet Identification (b-tagging)

- B-hadrons are long-lived – search for displaced vertices
- Construct event-by-event primary within beamspot (10-32 μm)
- Fit displaced tracks and cut on $L_{xy}$ significance ($\sigma \sim 200$ μm)
- Calibrate performance from data (low-$p_T$ lepton samples)

- Charm hadrons have similar tag behavior but lower efficiency

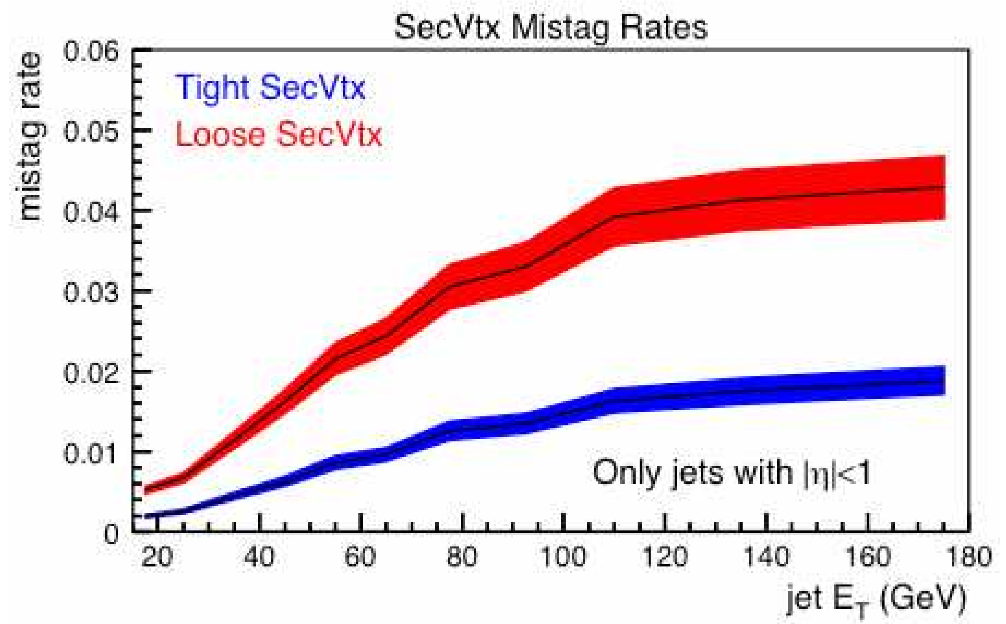


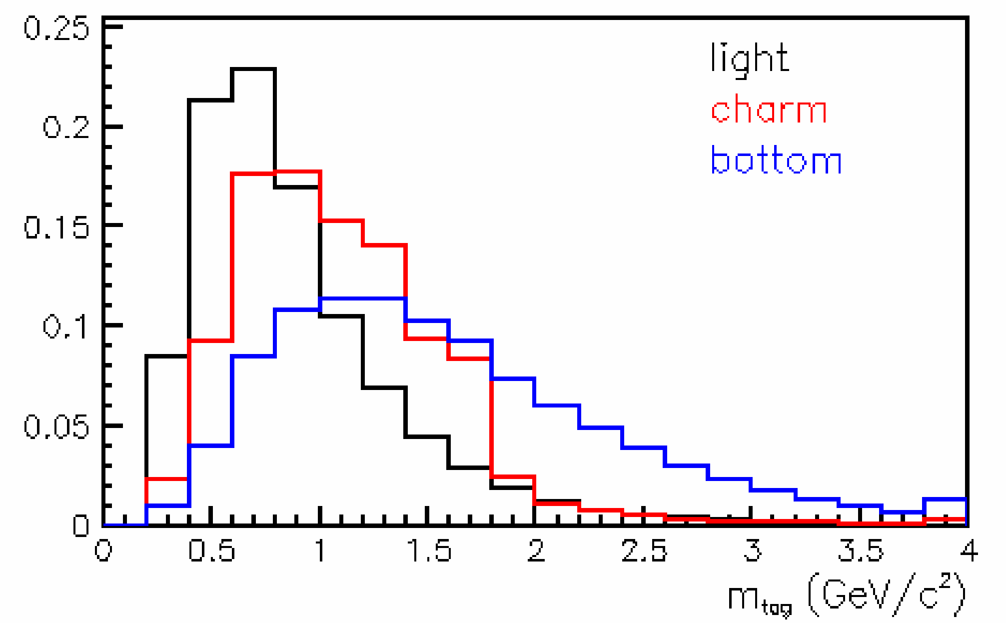



# Fake B-Tags (mistags)

- Fake tags of light-flavor jets can result from poorly-reconstructed tracks or interactions with the detector material
- More tracks in a jet above 1 GeV/c means more chances for something to go wrong – fake rate rises dramatically with jet $E_T$




- Can use "tag mass" to deduce the flavor composition of a sample of tagged jets
  - Mass of the tracks forming the secondary vertex
  - B-hadrons are heavy → will have higher $m_{tag}$ spectrum than charm or light jet fakes



# 3b Channel Roadmap

- Trigger is two jets and two displaced tracks (no matching)
  - Events are fairly biased even after offline selection
  - Data-driven backgrounds are the natural way to handle this

- Background is QCD multijet production of all possible jet flavors
  - We consider b-jets, c-jets, and “q”-jets (q = light quark or gluon)

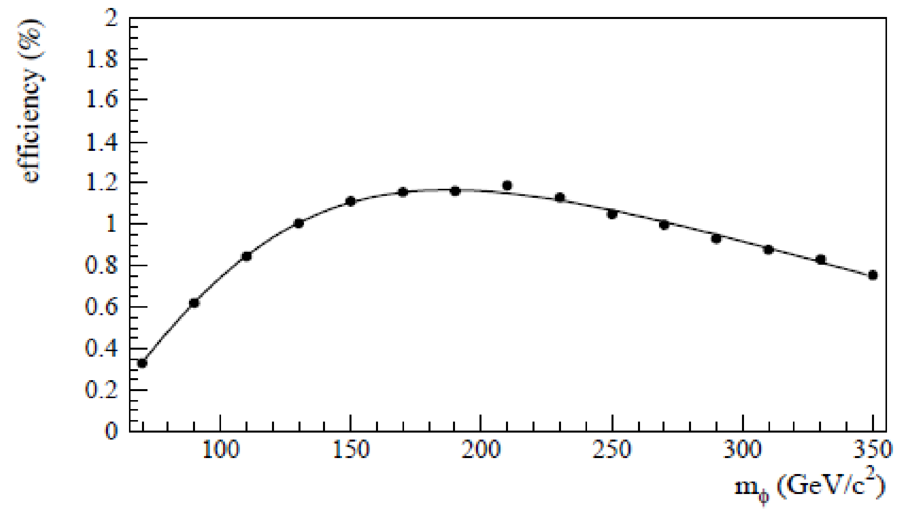
- Derive estimates for each flavor combination from the data
  - Use Pythia to check for bias

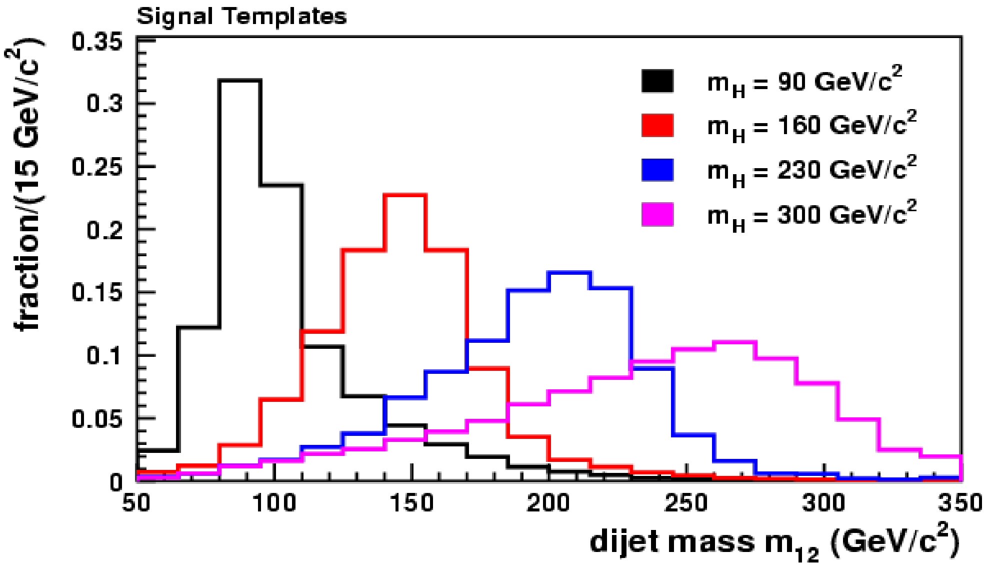
- Look for an excess in the mass of the two leading jets ($m_{12}$)
- Use tag mass ($m_{tag}$) information to understand flavor composition

- Perform a two-dimensional fit to the data using these estimates
  - Tag mass information determines background composition
  - Look for Higgs in the $m_{12}$ distribution

# Selection and Acceptance

- Base selection is three of the four leading jets must have $E_T$>20 GeV, $|\eta|$<2, and be b-tagged
- Two lead jets must be b-tagged, have two matched trigger displaced tracks (SVT) between them
  - Either 1+1 or 2+0
- Third or fourth-leading jet must have standard b-tag
- Signal samples are PYTHIA gg→bbH, H→bb
  - Require $p_T$>15 GeV/c, $|\eta|$<2 for at least one of the associated b-quarks
- Discriminating variable is mass of two leading tagged jets $m_{12}$



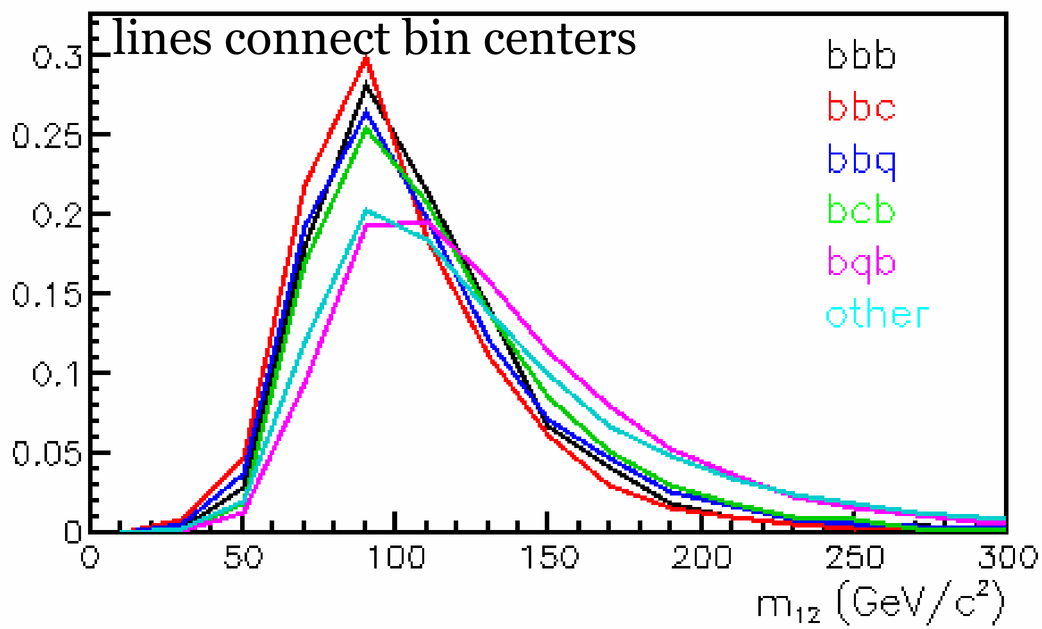
# Background Model

- Essentially ~100% of background is QCD multijet production
- Most of it is in five components – two b's plus b/c/q

- Fractions are from generator-level Pythia samples
  - Do not want to rely on these cross sections
  - Will fit for them in the data

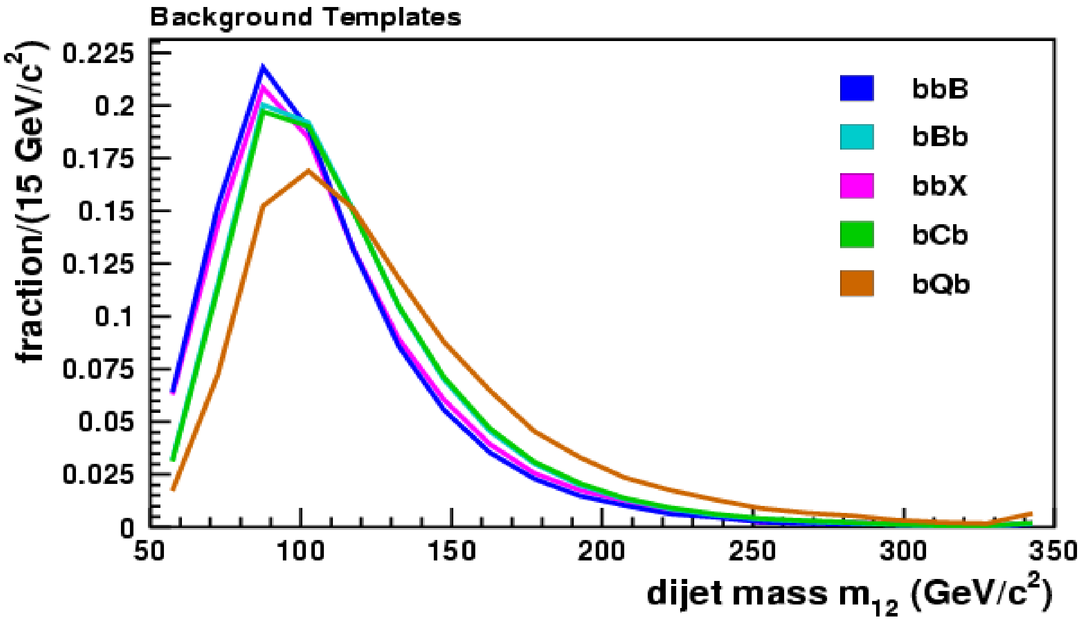
- Shape differences are due to both physics and detector effects
  - Biases from b-tagging
  - Kinematics of heavy quark production processes

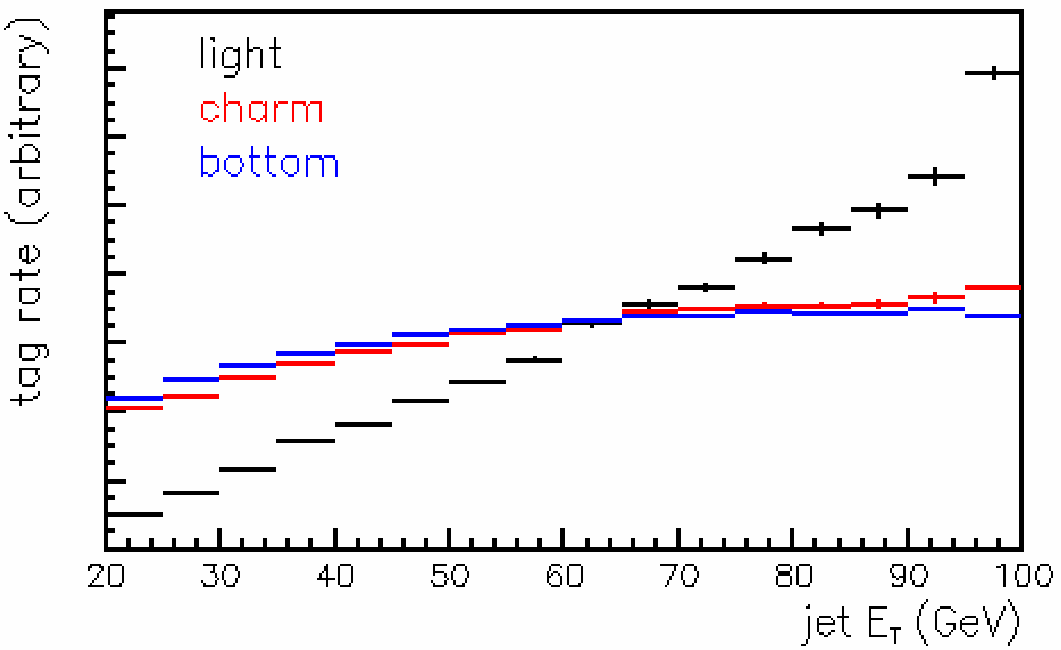
| jet flavor (12)3 | fraction |
|---|---|
| bbb | 0.37 |
| bbc | 0.05 |
| bbq | 0.10 |
| bcb | 0.12 |
| bqb | 0.29 |
| other | 0.07 |



# Background Strategy – $m_{12}$

- Use three-jet events with two jets b-tagged as starting point
- For bbq and bqb it's straightforward
  - Weight events by mistag probability of untagged jet
  - Physics from data, we add tagging bias
- Tag probabilities: functions of jet $E_T$ and number of tracks, derived from large simulated jet samples
- For bbc and bcb, PYTHIA studies show that 'c' events have same kinematics as 'q'
  - Weight by probability to b-tag untagged jet as if it were a charm jet
- Comparing bCb vs bQb we see the effect of the b-tagging bias
- Bias from third-jet b-tag not that important for $m_{12}$ → combine bbQ/bbC into 'bbX'



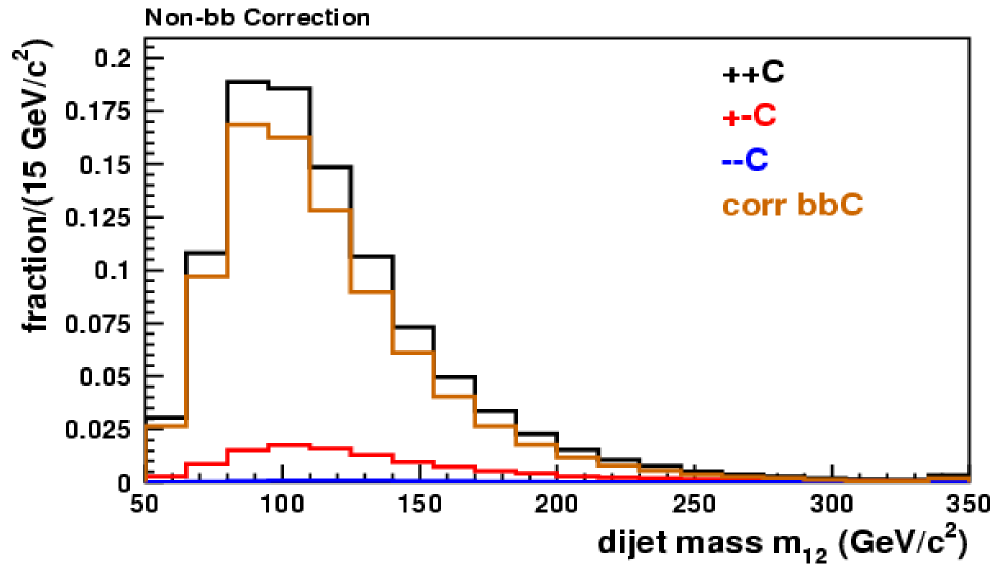
# Non-bb Subtraction

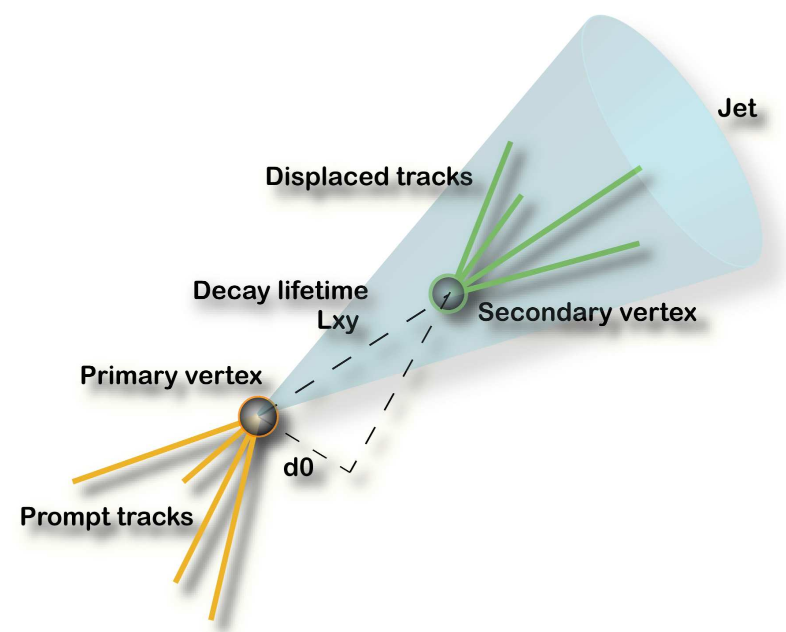
- Using double-tagged events as an estimate of bb+b/c/q events
- Double-tags contain some non-bb component

- Subtract using 'negative' tags
  - Vertex which is on the opposite side of the beamline compared to the jet direction
  - Often used as an estimate for light-flavor fake tags

- Perform same weighting procedure on events with one or two negative tags, and compute
  - $N_{bb} = N_{++} - \lambda N_{+-} + \lambda^2 N_{--}$

- The factor $\lambda$ is the ratio of POS/NEG fake tag rates (1.4)



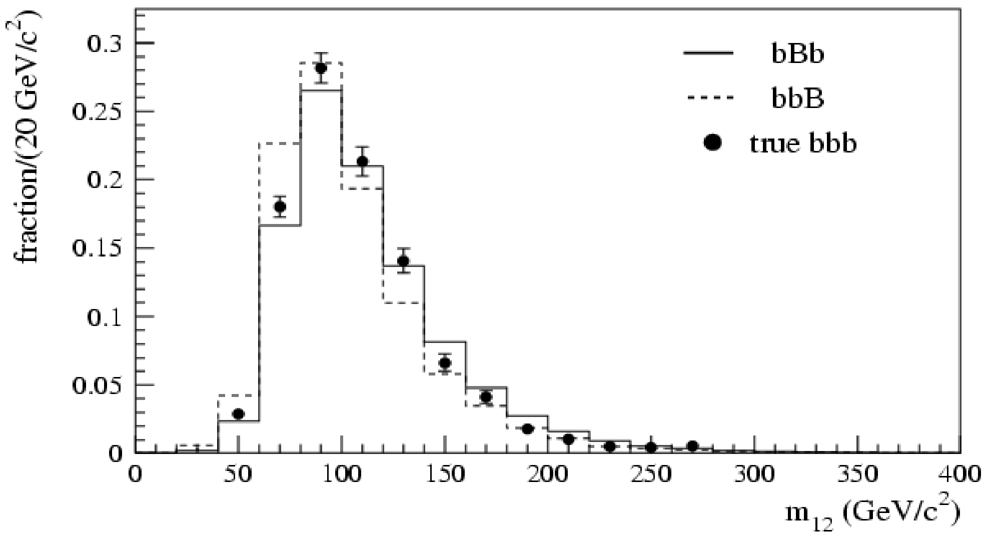
# What about bbb?

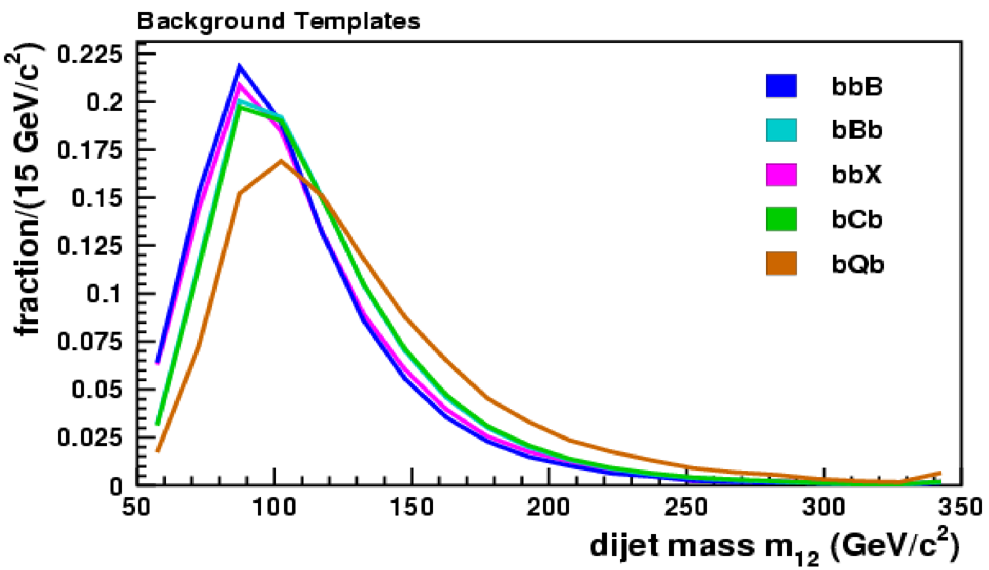
- Here we have two options for the starting point
  - Two lead jets b-tagged (bbB)
  - Third jet and either of two lead jets b-tagged (bBb)
- Which one is correct?

- Answer: neither

- The true bbb background has contributions from events where $m_{12}$ reflects a single bb production pair (bbB), and where the two jets are from different production pairs (bBb)

- We include both cases in the fit



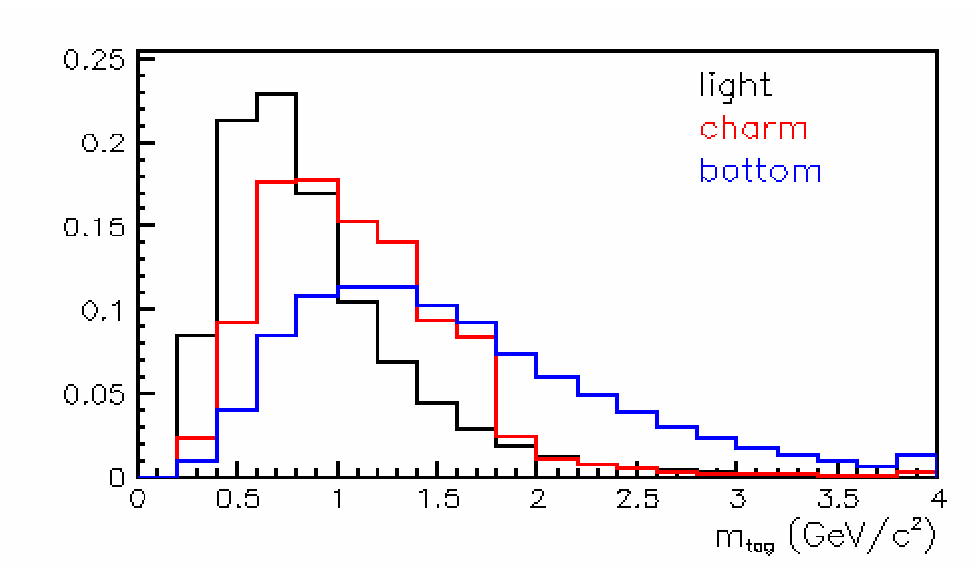
# Second Template Axis

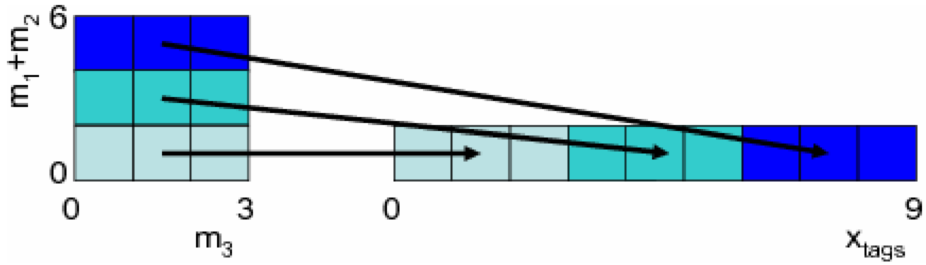
- Split events by flavor
- Characteristic $m_{12}$ spectra

- Second variable to help separate backgrounds from each other, and Higgs+bbb from ones with c/q

- Important categories are:
  - bb + b : bbb, Higgs
  - bb + X : bbc, bbq
  - bX + b : bcb, bqb
  - Naturally breaks into $m_1+m_2$ and $m_3$

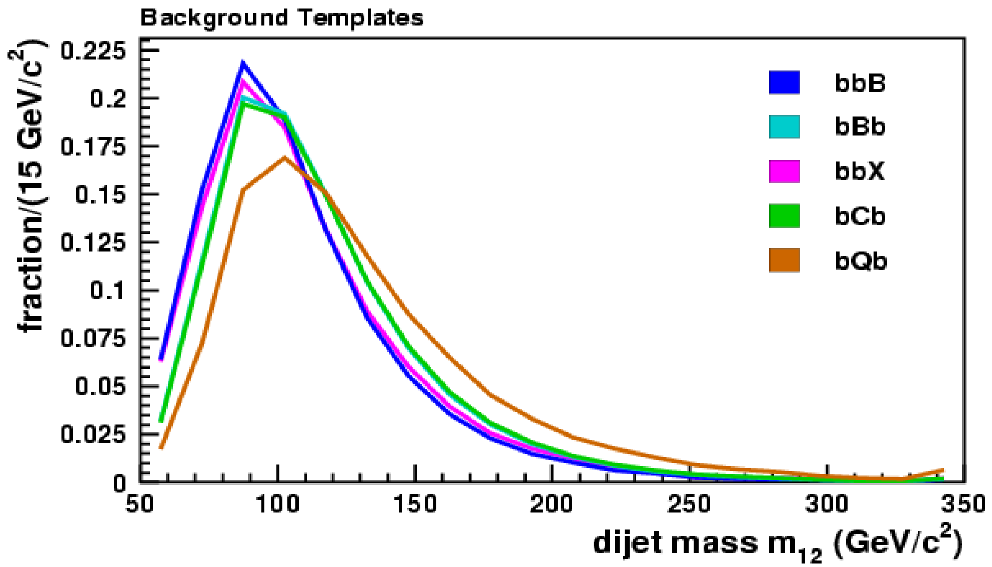
- Pack into 1D so overall templates are only 2D (technical reasons)
- Unstack 3x3 histogram into a 9-bin 1D histogram – "$x_{tags}$"

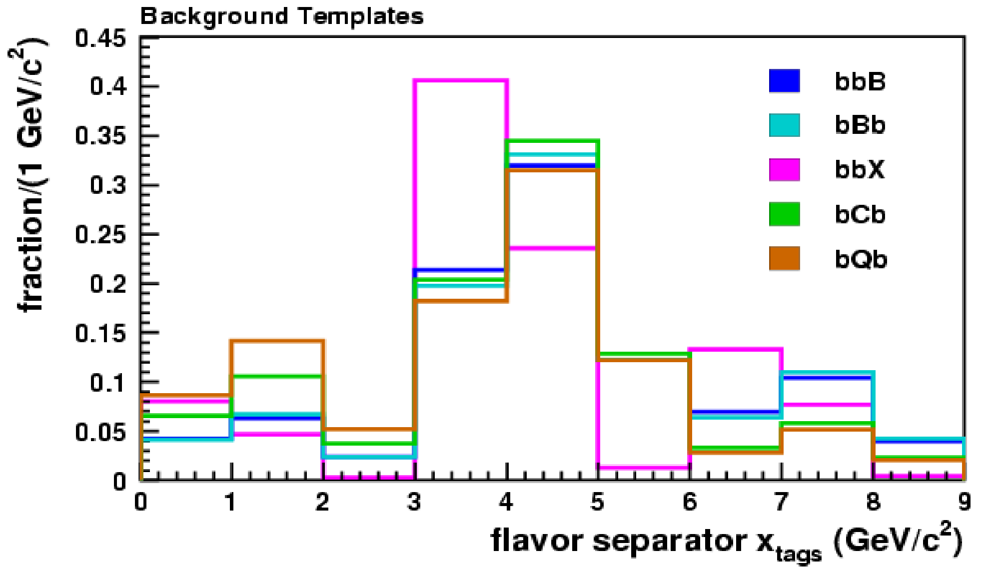


# Fit Templates

- The bbX events can be separated by third tag mass in $x_{tags}$
- Two lead jet tag masses separate bbB,bBb from bCb,bQb
- Separation of bCb vs bQb and bbB vs bBb vs Higgs through $m_{12}$
- Templates are actually 2D histograms in both $m_{12}$ and $x_{tags}$
  - Fit itself is also 2D
  - Only show projections for clarity



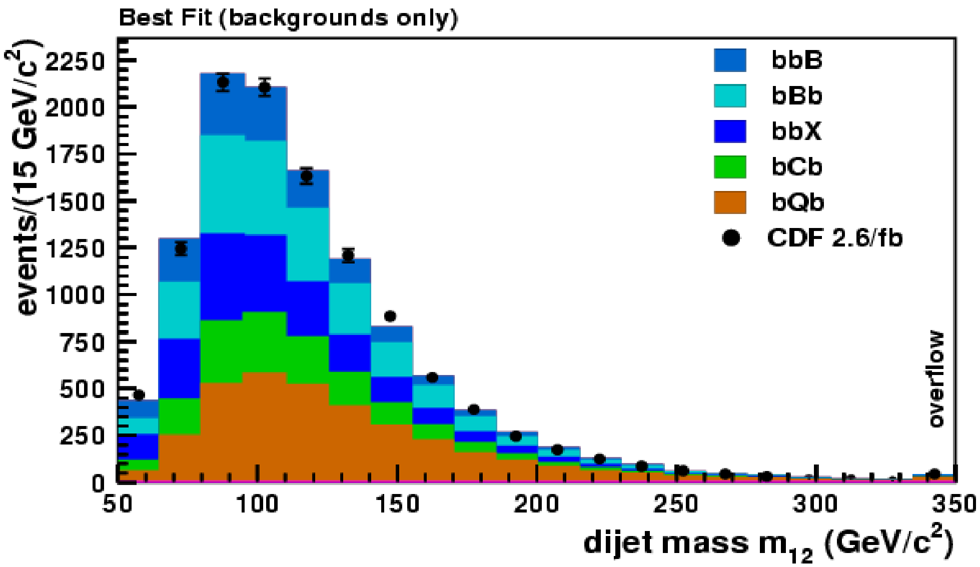



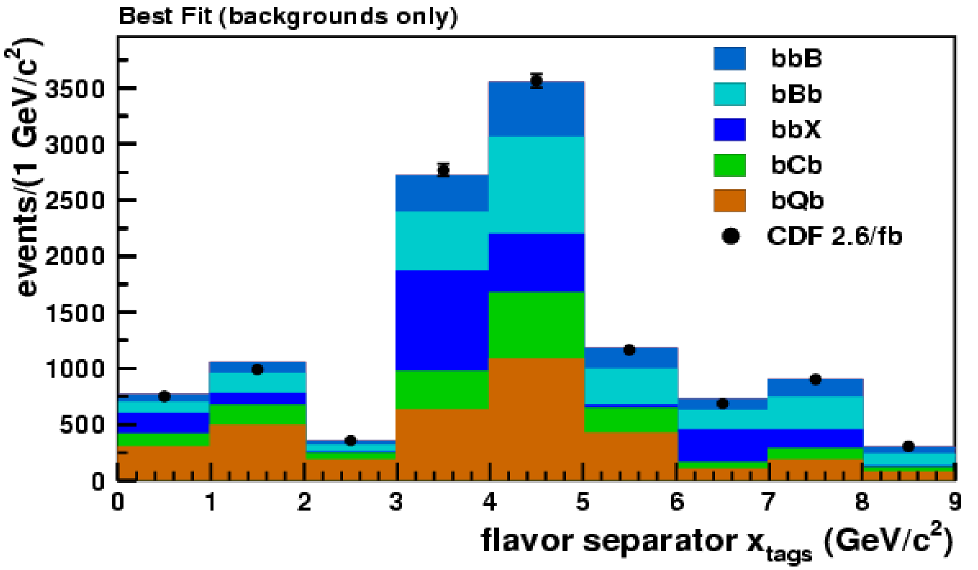
# Background-Only Fit

- Total of 11490 events in the data
- Simple template fit with MINUIT, no systematics
- Results in 2.6/fb

| component | estimate | $N_{fit}$ |
|---|---|---|
| $bbB$ | 1300 | 1520±540 |
| $bBb$ | 2950 | 2620±550 |
| $bbX = bbQ + bbC$ | 1350+640 = 1990 | 2210±160 |
| $bCb$ | 1380 | 1710±630 |
| $bQb$ | 3480 | 3430±390 |

- Fit $\chi^2$/dof = 1.14



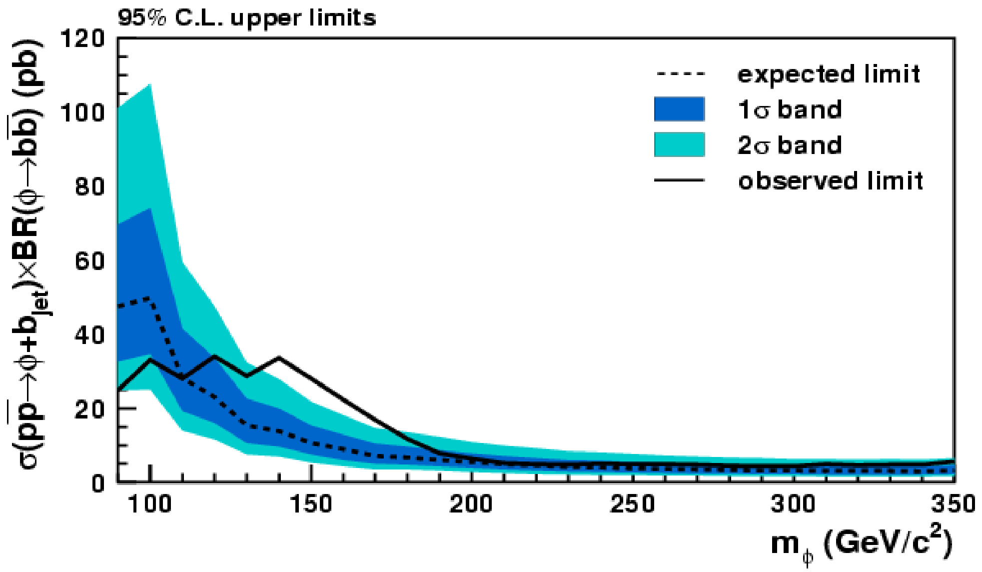
# Limits and Systematics

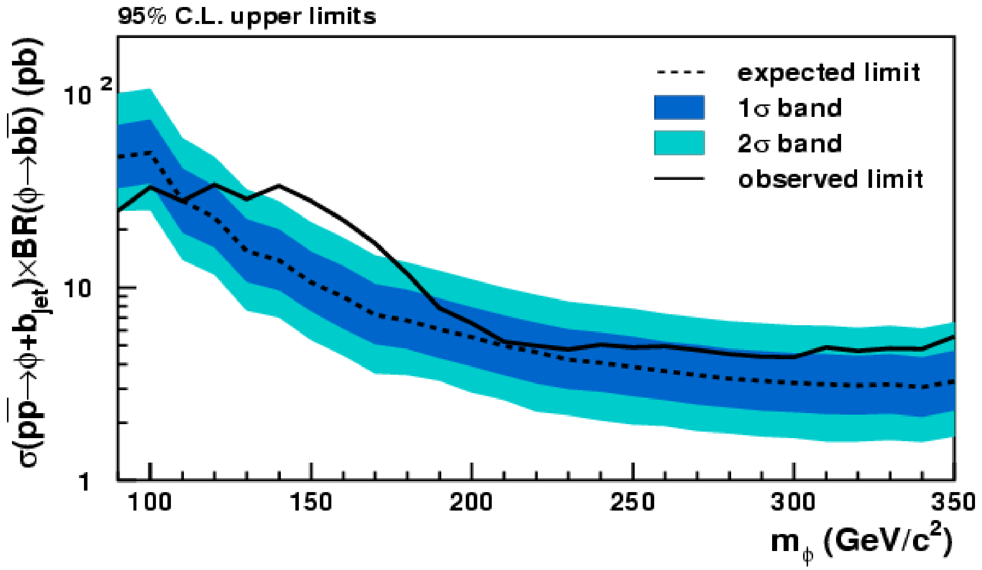
- Modified-frequentist ($CL_S$) limit calculator based on mclimit_csm
  - Allows inclusion of systematic uncertainties
  - Shape uncertainties on templates through histogram interpolation

| source | variation | applies to | type |
|---|---|---|---|
| luminosity | $\pm 6\%$ | signal | rate |
| Monte Carlo statistics | $\pm 2\%$ | signal | rate |
| selection efficiency | $\pm 5\%$ per jet | signal | rate |
| PDFs | $^{+3.5}_{-4.5}\%$ | signal | rate |
| jet energy scale | $\pm 4.5\%$ | signal | rate/shape |
| $b/c\ m_{tag}$ | 3% | signal/backgrounds | shape |
| mistag $m_{tag}$ | 3% | backgrounds | shape |
| mistag asymmetry factor $\lambda$ | $1.4 \pm 0.2$ | backgrounds | rate/shape |
| heavy flavor fractions | $\pm 50\%$ | backgrounds | rate |

# σ x BR Limits

- Maximum deviation from expectation at 150 GeV/c²
- Including the trials factor, 1-$CL_b$ = 2.5% (1.9σ)
- For just the 150 point alone, 1-$CL_b$ = 0.23% (2.8σ)
- Corresponds to σ x BR ~ 15 pb
- Note: these limits are for a resonance much narrower than the experimental resolution!
  - SM Higgs, new scalars, etc
- MSSM Higgs in high-tanβ scenarios not generally narrow



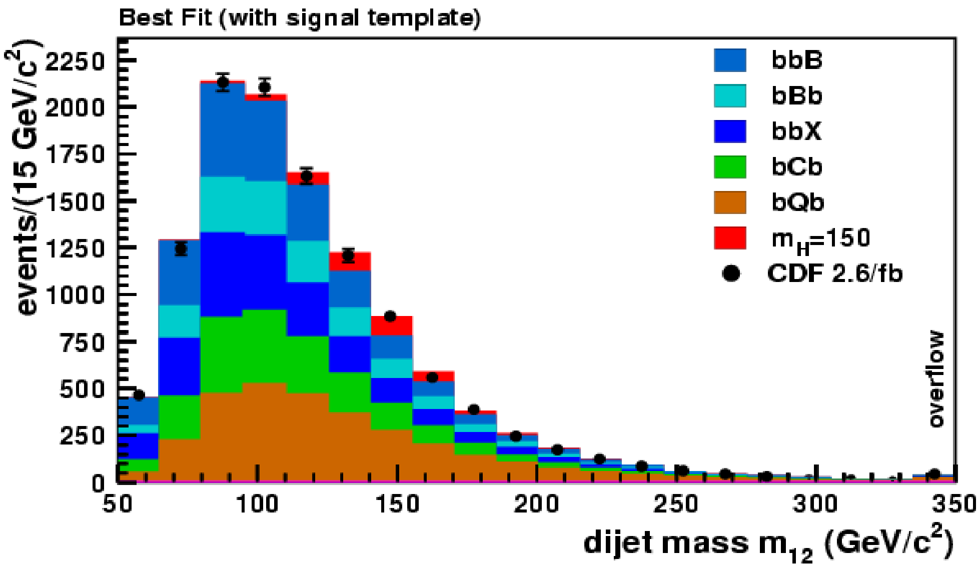
# Fit Including Signal

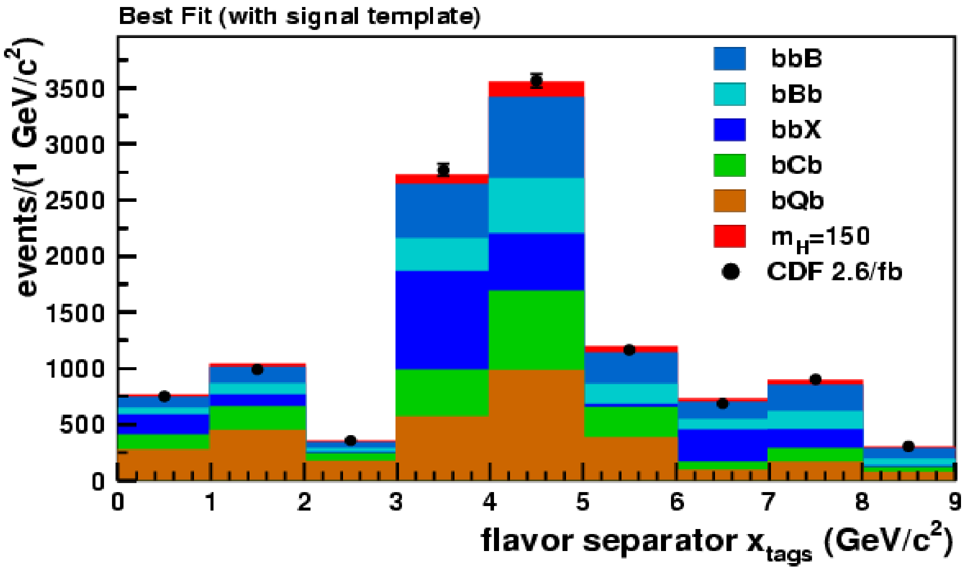
- Repeat simple fit, using backgrounds plus a Higgs signal component ($m_H$ = 150 GeV/c²)

- Results in 2.6/fb

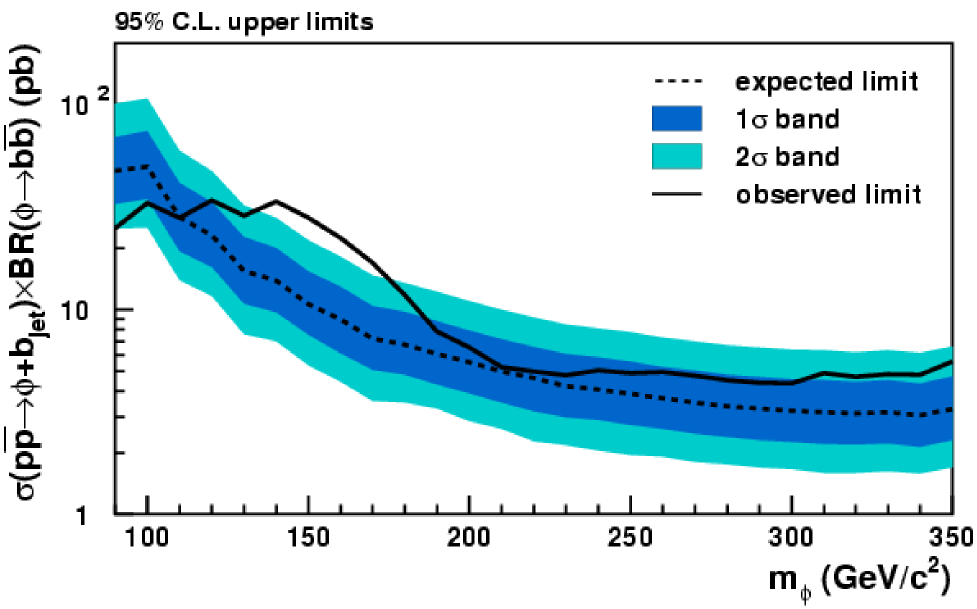
| component | $N_{fit}$ |
|---|---|
| $bbB$ | 2280±600 |
| $bBb$ | 1490±670 |
| $bbX$ | 2150±160 |
| $bCb$ | 2050±630 |
| $bQb$ | 3100±400 |
| Higgs | 420±130 |

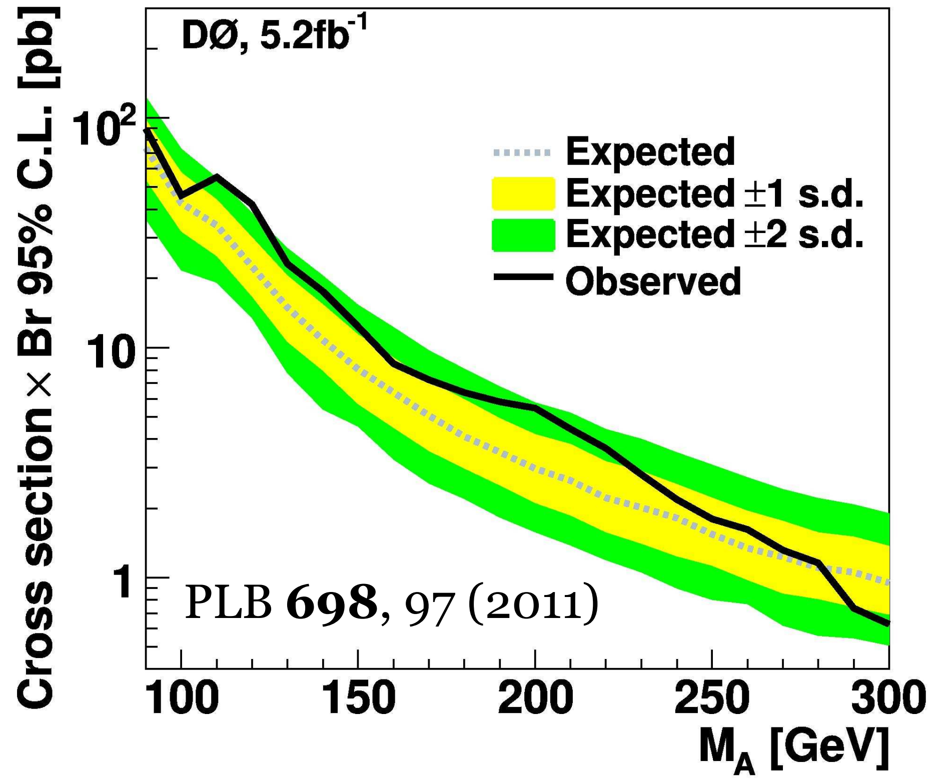
- Fit $\chi^2$/dof = 1.06



# Current σ x BR Constraints




- D0 have a similar analysis (J. Hays W&C 11/12/2010)
  - No excess at 150 GeV/c²
  - Different cross section definitions (D0's is a factor ~2 larger)

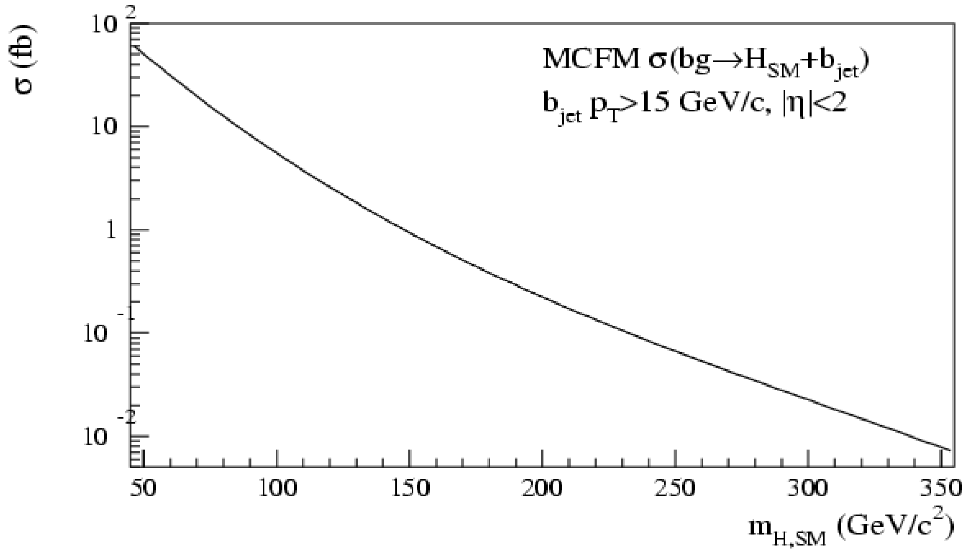
- Work is in progress to combine the two results

# MSSM $\tan\beta$ Limits

- At tree level,
  $\sigma$ x BR = $2\sigma_{SM}\tan^2\beta$ x 90%

- Get $\sigma_{SM}$ from MCFM
- Factor of 2 from h/H degeneracy
- BR = 90%

- For $\tan\beta$ limits, must include uncertainties on the predicted cross section
  - PDFs:
    - was 4% on the acceptance
    - change to uncertainty on the total event yield
    - 8%(23%) for $m_H$=90(210)
  - Additional variation for scale dependence of NLO calculation



# Scenario Dependence

- Higgs properties are largely, but not completely, determined by $m_A$ and tanβ
- Loop corrections introduce dependence on other SUSY parameters
  - M. Carena *et al.*, Eur.Phys.J. C45 (2006) 797-814 (hep-ph/0511023)

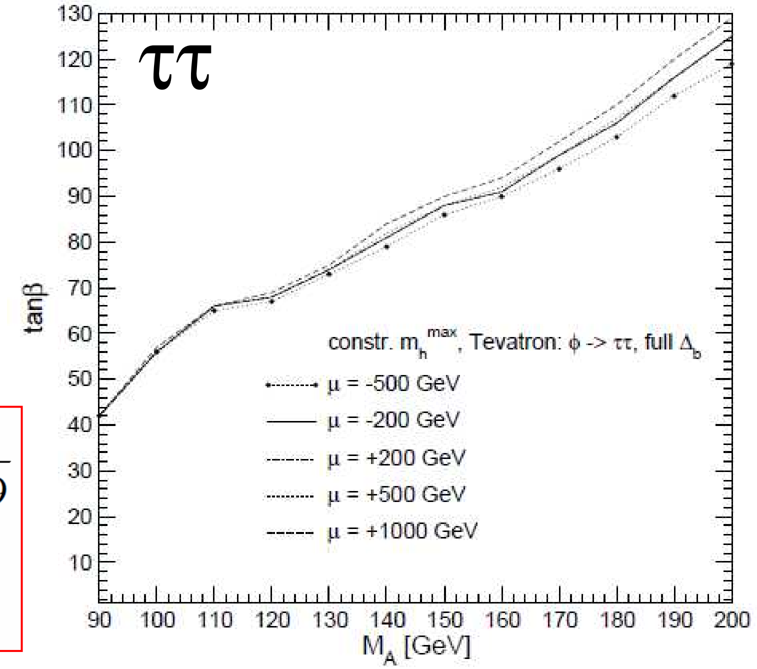
$$\sigma(b\bar{b} \to A) \times BR(A \to b\bar{b}) \cong \sigma(b\bar{b} \to A)_{SM} \times \frac{\tan\beta^2}{(1+\Delta_b)^2} \times \frac{9}{(1+\Delta_b)^2 + 9}$$

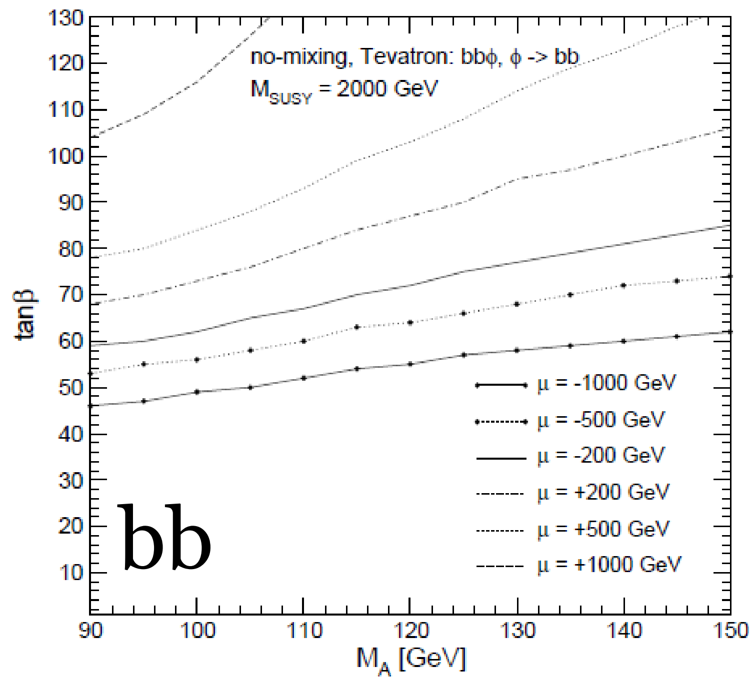
$$\sigma(b\bar{b}, gg \to A) \times BR(A \to \tau\tau) \cong \sigma(b\bar{b}, gg \to A)_{SM} \times \frac{\tan\beta^2}{(1+\Delta_b)^2 + 9}$$

- $\Delta_b$ is a function of SUSY parameters
  - $\Delta_b$ proportional to μ x tanβ (sign of μ matters)

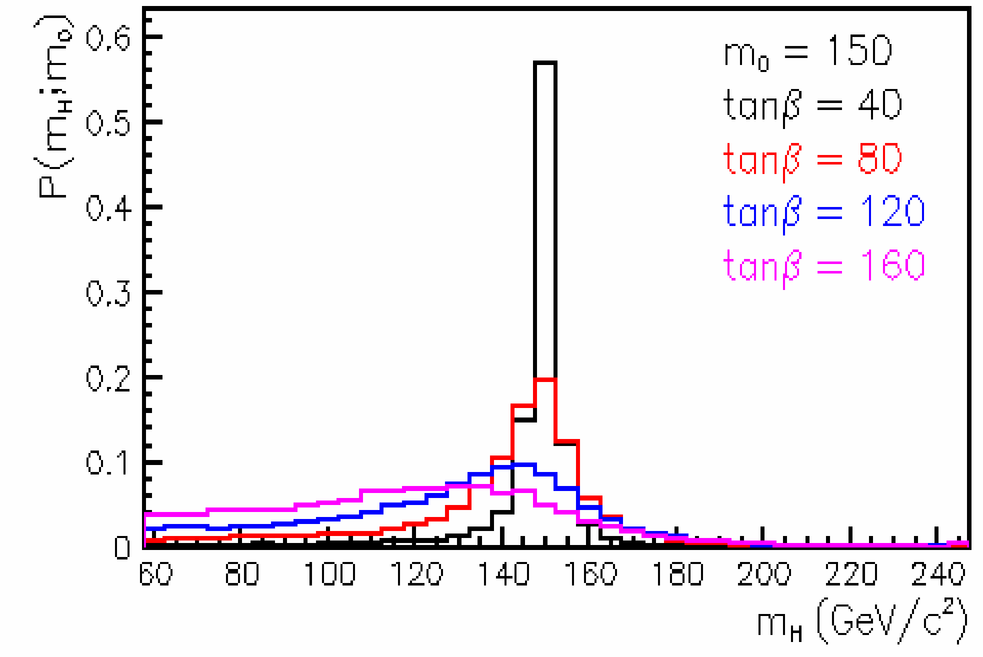
- Use two “benchmark” scenarios

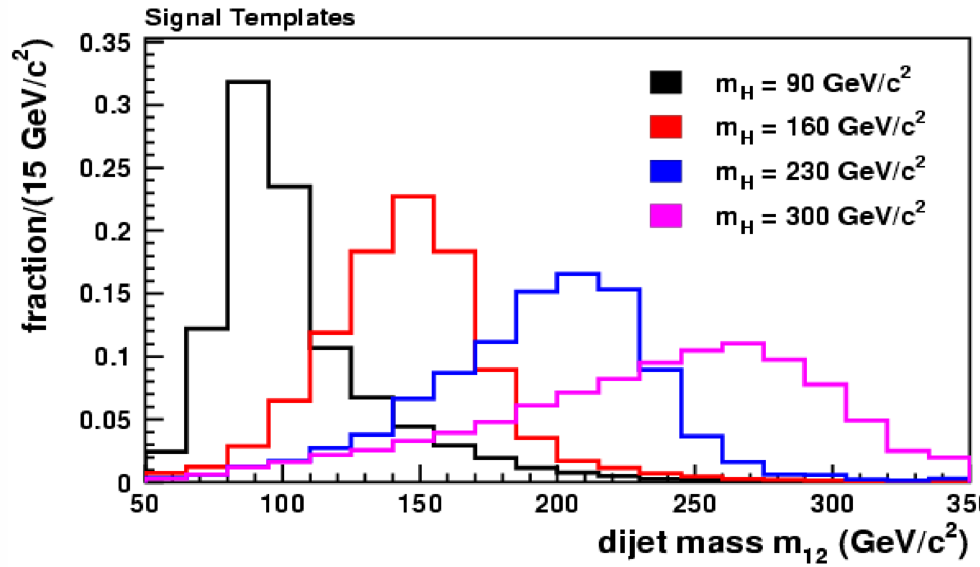
- For tanβ = 50, μ = -200 GeV
  - $m_h^{max}$: $\Delta_b$ = -0.21
  - tree-level: $\Delta_b$ = 0

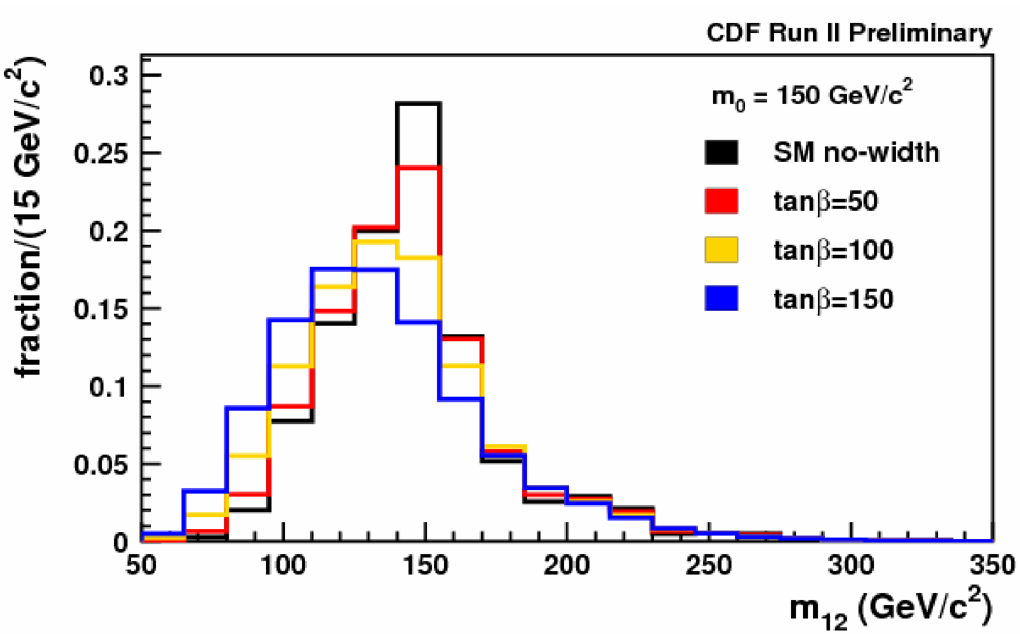


# Fit Templates with Width

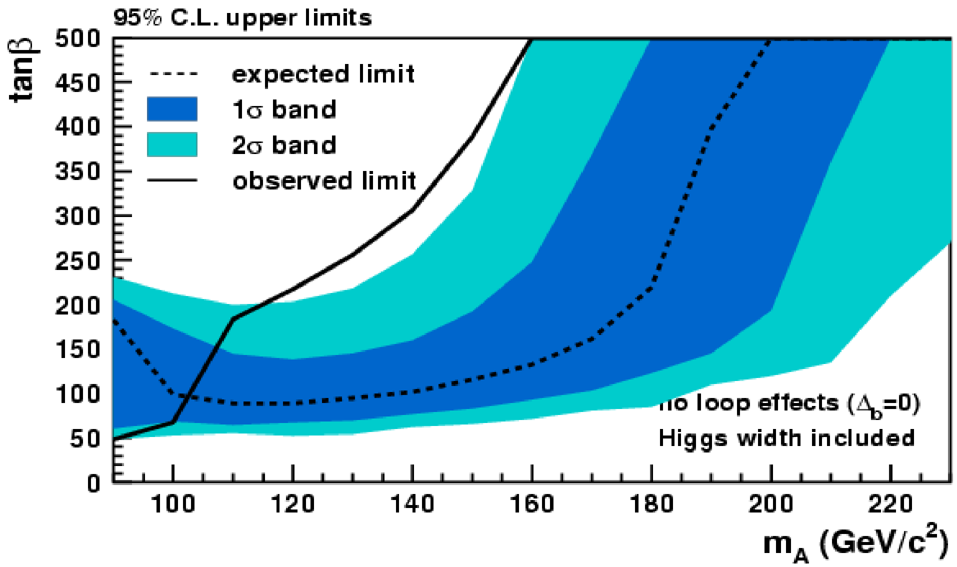


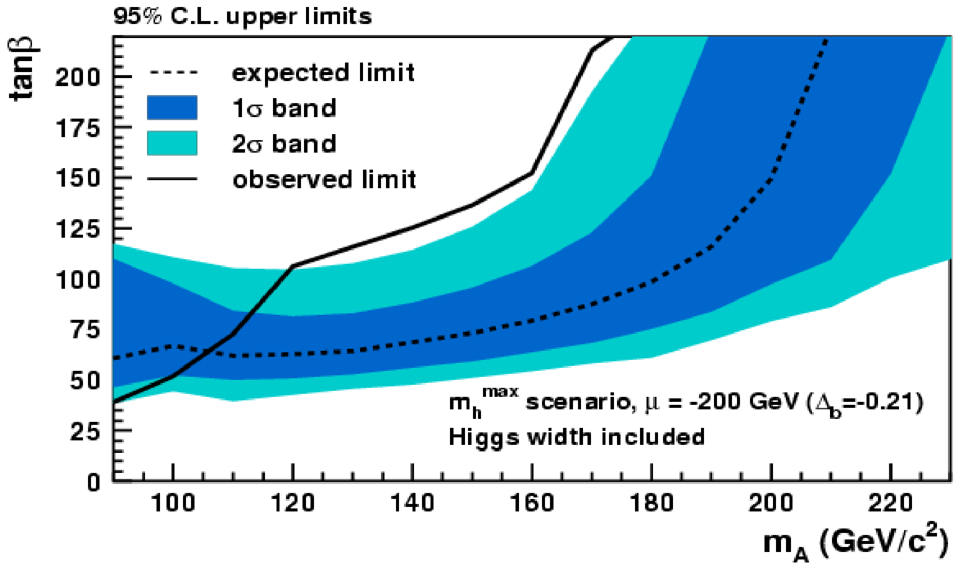



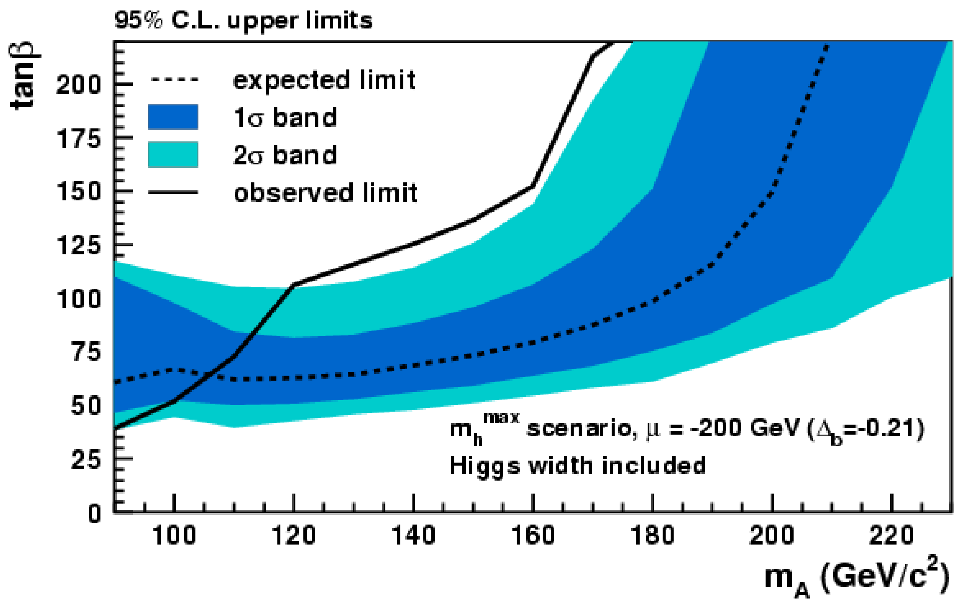

- Can't turn up cross section without increasing Higgs width
- Natural width eventually becomes comparable to experimental resolution
- Fit templates broaden and shift towards lower $m_{12}$ because of falling cross section

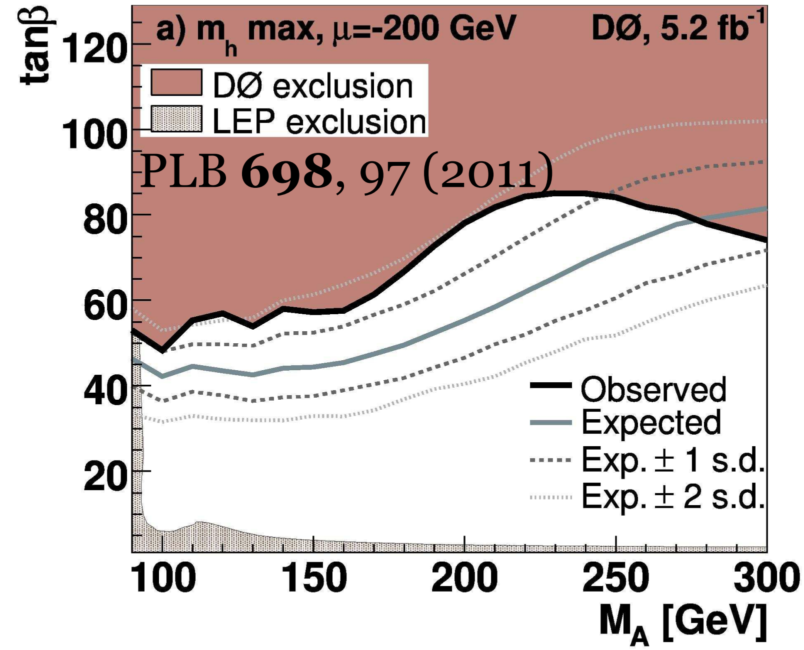
- Upper plot is for the tree-level case ($\Delta_b$=0) and with the width effect included
- Lower plot is for $m_h^{max}$ scenario with negative μ and the width effect
  - Enhanced signal production improves the limits
  - Effect is greatest at large tanβ because $\Delta_b \sim \tan\beta$
- With enough data the width effect would eventually die out
  - Sensitivity to smaller cross sections means we probe narrower Higgs regime
  - Will see what we get with the full Run II sample

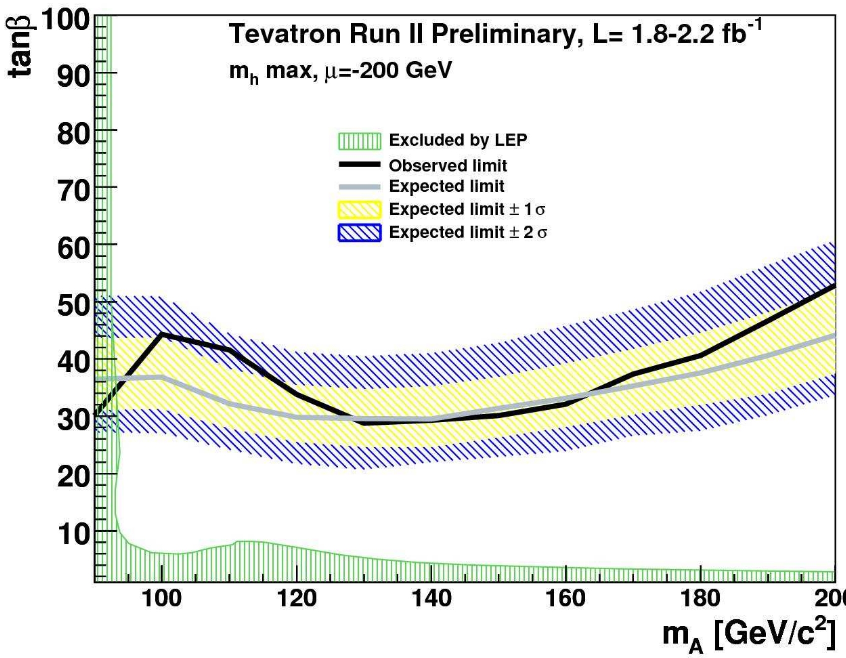


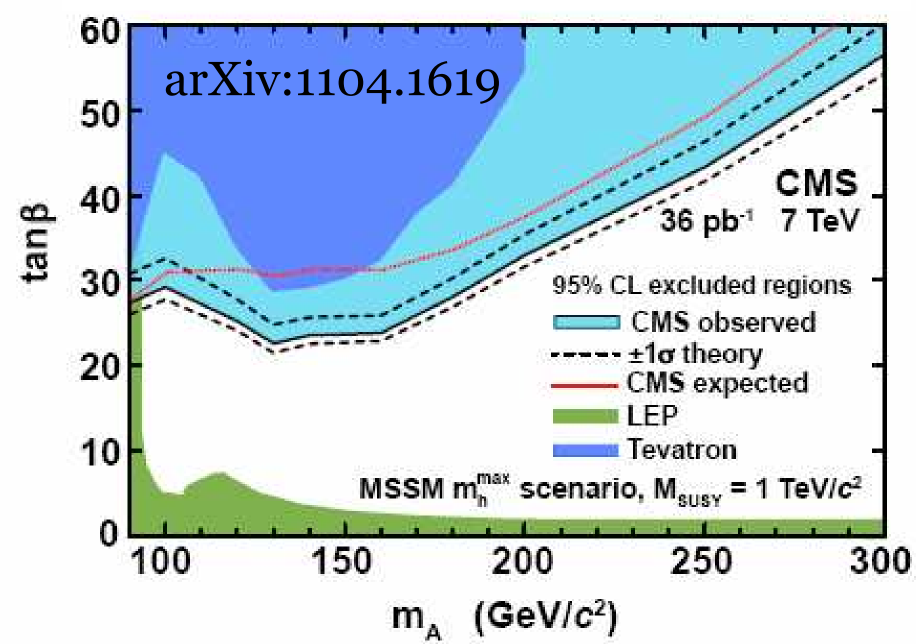



# Current MSSM Constraints



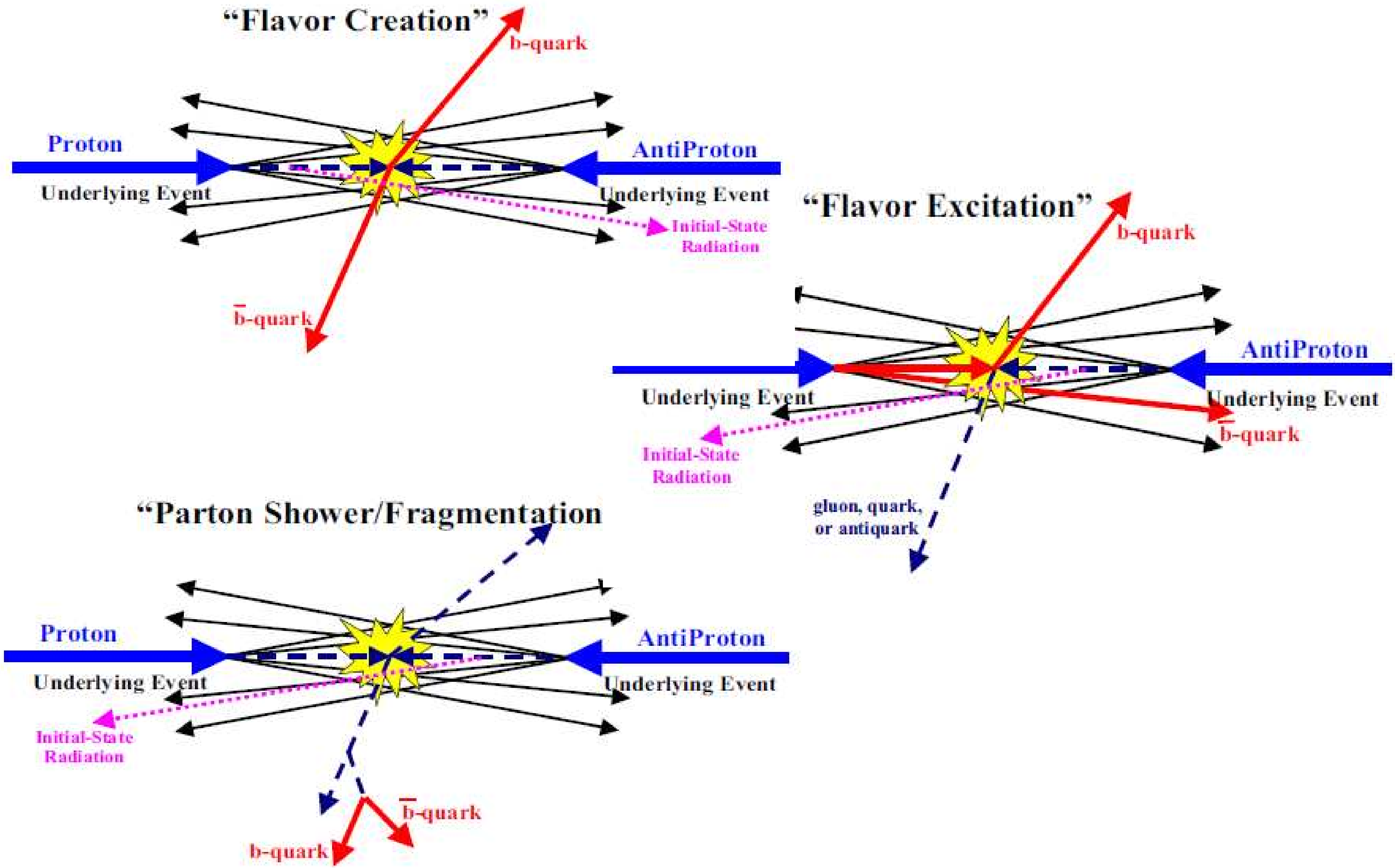



# Summary and Outlook

- CDF has updated MSSM Higgs results in the 3b channel
  - Submitted for publication in PRD (arXiv:1106:4782)
- Analyzed all data taken with the original b-jets trigger (2.6/fb)
  - Still a considerable amount remaining
  - Will update with full Run II data sample
- No significant excess observed, ~2$\sigma$ bump at 150 GeV/$c^2$
- Analysis is adaptable to other signal models besides MSSM Higgs
  - Ex: color octet scalars (B. Dobrescu's W&C 10/15/2010)
- Paper, plots, etc available at
  - http://www-cdf.fnal.gov/physics/new/hdg/Results.html

# Backup Material

# QCD Heavy Flavor Production




thanks to Rick Field